Biochemical Studies On *Methanobrevibacter smithii* RNase P – A Model For Mesophilic Type A Archaeal RNase P And A Possible Anti-Obesity Target

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation "with honors research distinction" in the undergraduate colleges of The Ohio State University

By

Emily Wong

The Ohio State University

June 2011

Project Advisor: Venkat Gopalan, Department of Chemistry and Biochemistry

TABLE OF CONTENTS

ABBREVIATIONS

$A_{260/280}$	ratio of absorbance at 260 and 280 nM
DNA	deoxyribonucleic acid
DTT	dithiothreitol
Eco	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
HPLC	high performance liquid chromatography
IPTG	isopropyl β–D-1-thiogalactopyranoside
IVT	*in vitro* transcription
kDa	kilodalton
LB	Luria-Bertani
Mja	*Methanocaldococcus jannaschii*
Mma	*Methanococcus maripaludis*
Msm	*Methanobrevibacter smithii*
Mth	*Methanothermobacter thermoautotrophicus*
MTO	multiple turnover
NaOAc	sodium acetate
NH_4OAc	ammonium acetate
OD_{600}	optical density at 600 nM
PCR	polymerase chain reaction
PEI	polyethylene imine
Pfu	*Pyrococcus furiosus*
Pho	*Pyrococcus horikoshii*
PMSF	phenylmethylsulfonyl fluoride
pre-$tRNA^{Tyr}$	precursor $tRNA^{Tyr}$
RNA	ribonucleic acid
RNase P	ribonuclease P
RNP	ribonucleoprotein

RPP	RNase P Protein
RPR	RNase P RNA
SDS	sodium dodecyl sulfate
sm-FRET	single molecule-Fluorescent Resonance Energy Transfer
STO	single turnover
TFA	trifluoroacetic acid
Tris	tris(hydroxymethyl)aminomethane

ABSTRACT

RNase P, a universal and essential catalytic ribonucleoprotein responsible for cleaving the 5'-leader of precursor tRNA transcripts, has been successfully reconstituted from several different archaea. Although both types of archaeal RNase P—types A and M, which are similar to bacterial and eukaryotic RNase P, respectively—have been reconstituted, all archaeal type A holoenzymes studied to date are thermophilic in origin and require high temperatures ($\geq 55^{o}$C) for optimal pre-tRNA 5'-processing activity. The availability of a mesophilic archaeal RNase P, such as the one from *Methanobrevibacter smithii*, however, will permit single molecule-FRET and rapid quench flow kinetic studies, thus far not possible due to limitations imposed by the optimal temperature for thermophilic archaeal RNase P activity. Other payoffs from studies on *M. smithii* RNase P are likely. Since *M. smithii*, the predominant archaeon in the human gut, is a methanogen whose syntrophic relationship with saccharolytic bacteria permits increased fermentation of sugars in the digestive process, and ultimately a greater caloric absorption from food, it is a potential anti-obesity target. As the subunit makeup of archaeal RNase P differs from its eukaryotic and bacterial relatives, it might be possible to design a specific inhibitor of *M. smithii* RNase P without harming the animal host or commensal bacteria. Towards the goal of biochemical characterization of the *M. smithii* RNase P holoenzyme, cloning, overexpression, and purification of the protein subunits has been successfully achieved, and the reconstituted holoenzyme shows activity at 37^{o}C. Kinetic studies of the *M. smithii* RNase P holoenzyme are currently in progress, and preliminary data suggest that the *M. smithii* RNase P does in fact share many characteristics with other archaeal type A RNase P holoenzymes and might prove to be a useful experimental paradigm.

ACKNOWLEDGEMENTS

I would like to thank my adviser, Dr. Venkat Gopalan, and the members of his laboratory for their knowledge and expertise not only at the bench but also away from it, and for their extreme patience in imparting both to me. I am grateful to Dr. Mark Foster and his laboratory members for sharing their HPLC and reagents. I would also like to thank my committee members, Dr. Juan D. Alfonzo and Dr. Dmitri Kudryashov, for their suggestions and guidance. Last but not least, I would like to thank the Beckman Foundation and the Pelotonia Fellowship Program for generously funding my research experience over the past two years.

INTRODUCTION

Although most introductory-level biology students are taught that enzymes are typically proteins that catalyze reactions within the cell, some of the most fundamental reactions of life are actually the realm of catalytic ribonucleoproteins (RNPs). These complexes are comprised of both ribonucleic acid (RNA) and protein. While the catalytic function may reside in either protein or RNA component, cooperation between both types of subunits is critical for the overall biological function of the complex. Telomerase, for example, preserves the integrity of the linear ends of the eukaryotic chromosome by using an RNA molecule as a template for reverse transcription by a protein subunit.[1] However, there are also many examples of the opposite situation in which the RNA is the catalyst rather than an information carrier. The ribosome, a more universal RNP, is tasked with the polymerization of individual amino acids into functional polypeptides, and the spliceosome is responsible for the excision of introns from pre-mRNA—in both these instances, catalysis is believed to reside with the RNA portion of the ribosomal and spliceosomal RNP complexes.[2,3] Regardless of the component in which the enzymatic activity resides, all of these complexes perform reactions that are necessary for viability. It is also notable that, unlike many other enzymes, there is little redundancy in the function of RNPs as borne out by the absence of cellular backup mechanisms capable of rescuing a deficiency in the function of these vital RNPs. The general study of catalytic RNPs is thus essential to gain a deeper understanding of life, both in its present form as well as from an evolutionary context in terms of the transition from a putative RNA world to enzymes made up of RNPs and proteins. The focus of this dissertation is an ancient catalytic RNP called ribonuclease P (RNase P).

A universal and highly conserved RNP, RNase P is the only known entity responsible for catalyzing the removal of the 5' leader from precursor transfer RNAs (pre-tRNAs).[4] It is therefore necessary for viability. Like the ribosome and spliceosome, the magnesium-dependent 5'-tRNA-processing reaction is catalyzed by the RNA component of the RNP holoenzyme, and thus all RNase P holoenzymes contain an essential RNA subunit.[5-9] (Eukaryotic mitochondrial and chloroplast RNase P,

believed to be wholly proteinaceous and devoid of RNA, are not discussed here.[10,11]) Although the ancestral (bacterial) RPR contains two independently folding domains (S domain, for substrate specificity, and C domain, for catalytic) the general structure of the RNase P RNA (RPR) varies between the three domains of life, as does the number of associated RNase P proteins (RPPs) (Fig. 1).[12,13] Bacterial RNase P has one RPP, archaeal has up to five, and eukaryotic has up to ten.[14-16] The variability in protein composition is counter-intuitive given that catalysis resides with the RPR rather than the RPPs. Thus, current research focuses on both the mechanism of RNase P catalysis and the role of the different RPPs in aiding RPR function.[17] Although such questions are most directly answered by structural studies in combination with characterization of the effects of the various RPPs on RNase P holoenzyme activity, of the two domains—bacterial and eukaryotic—which show the most evolutionary divergence, the eukaryotic RNase P holoenzyme has yet to be successfully reconstituted *in vitro*. This limitation has precluded detailed biochemical characterization and delineation of the intimate functional collaboration between a catalytic RNA and multiple protein subunits.

Because of the general intractability of eukaryotic RNase P to biochemical studies, many researchers currently use archaeal RNase P as a surrogate, inspired by the fact that the five archaeal RPPs (RPP21, RPP29, RPP30, POP5, and L7Ae) share homology with their eukaryotic counterparts.[18,19,20] Even within the archaeal domain, however, RNase P displays a variation in RPR structures (Fig. 2). The two main classifications are designated types A (for similarity to ancestral, or bacterial, RPR) and M (for the *Methanococcus* genera, whose RPR structure is representative of that category).[7] Both types have been successfully reconstituted and studied in some detail, heightening the expectation that some attributes shared with eukaryotic RNase P might well be delineated using the archaeal RNase P model. These studies entail the use of RNase P from thermophilic archaea, a source of highly thermostable enzymes. For example, the RNase P holoenzymes of the thermophilic archaea *Pyrococcus furiosus* (Pfu), *Pyrococcus horikoshii* (Pho), *Methanocaldococcus jannaschii* (Mja), and *Methanothermobacter thermoautotrophicus* (Mth) have all been successfully reconstituted *in vitro* and characterized.[21-24]

These investigations have provided valuable insights regarding archaeal RNase P structure and function. Studies on *Pfu* RNase P indicated that four of the five RPPs function as two heterodimers: RPP21 forms a binary complex with RPP29 (referred to hereafter as RPP21●RPP29) while RPP30 complexes with POP5 (RPP30●POP5) to activate the cognate RPR. NMR studies conducted on RPP21 and RPP29 showed significant chemical shifts upon the addition of the second protein, indicating a conformational change to accommodate an interacting partner.[19,20] Thematic parallels emerged from studies on RPP30 and POP5, collectively supporting the hypothesis that the two sets of RPPs form binary complexes both in the context of the holoenzyme and in solution.[19,20] The binary complexes were also found to footprint to different domains of the RPR, with RPP21●RPP29 localized to the S domain, and RPP30●POP5 to the C domain.[21,25] Furthermore, the high-resolution solution structure of some RPPs have been solved by either NMR and or X-ray crystallography, and the RNase P holoenzyme-pre-tRNA complex from the bacterium *Thermotoga maritima* was recently crystallized.[25,26,27] These findings have begun to provide a framework of the archaeal RNase P holoenzyme in terms of important RPR-RPP contacts, and even highlight differences with the bacterial cousin.

These structural studies have been complimented by investigations on the effects of the binary complexes on the functional activity of the *Pfu* holoenzyme which show that RPP30●POP5 is capable of increasing the k_{cat} by 100-fold, while RPP21●RPP29 is capable of decreasing the K_m by 16-fold.[23,28] Together, these findings have begun to define the mechanistic roles of the binary complexes based on their structure-function patterns. The ability of RPP21●RPP29 to bind to the S domain and the corresponding decrease in K_s provides evidence that the complex functions in facilitating substrate recognition, while the ability of RPP30●POP5 to bind the C domain and the corresponding increase in k_{cat} supports a role in facilitating chemical cleavage of the 5' leader once the substrate has been recognized and properly positioned. Given the above advances, it is now critical to understand the conformational dynamics during assembly of the tertiary structure of the RNP as a pre-requisite for elucidating the need for multiple RPPs in archaeal and eukaryotic RNase P.

One method of achieving that goal is to use single molecule-Fluorescent Resonance Energy Transfer (sm-FRET), which involves the strategic placement of two fluorophores, a donor and an acceptor, in the molecule under investigation. In this technique, the donor fluorophore is excited by radiation, and provided an acceptor is proximal, the energy will transfer to the acceptor fluorophore, resulting in the emission of light corresponding to the emission maximum of the acceptor even though the excitation wavelength corresponded to that of the donor.[29,30,31] By using sm-FRET to report on the tertiary structure of the RPR in solution, it should be possible to monitor the dynamic movement of the C and S domains as a consequence of RPR folding induced by metal ions and/or binding to RPPs. Other methods, like rapid quench flow kinetics and gel filtration, used to analyze enzymatic reactions at millisecond intervals and determine the stoichiometry of RPPs and RPR in the holoenzyme, respectively, are also expected to yield important information on the structure and function of the archaeal RNase P holoenzyme.

Unfortunately, these techniques are difficult to utilize with the set of model systems currently available (however, see note on *Mma* RNase P below). Of the systems previously mentioned, *Pfu*, *Pho*, and *Mth* are members of thermophilic type A RNase P, while *Mja* is classified as a thermophilic type M. These enzymes tend to display optimal activity at 55^{o}C or higher (some *in vitro* reconstituted and partially purified native holoenzymes have optimal activity up to 72^{o}C) which, while facilitating purification and reconstitution of robust activity, also tends to be above the functional temperature limits of sensitive laboratory equipment. As a result, the kinds of biochemical characterization pertinent to dissecting the assembly of RNase P are limited either by the temperature requirements of the laboratory equipment or of the thermophilic models. Thus there exists a need to complement the well-characterized thermophilic models with systems that operate at more moderate temperatures to employ certain techniques (like sm-FRET, rapid quench flow kinetics, and gel filtration) which are incompatible with high temperature conditions.

Last year, the successful reconstitution of the RNase P holoenzyme from *Methanococcus maripaludis* (Mma), a mesophilic archaeon with a type M RPR structure, was reported and this variant represents a mesophilic model for archaeal type M RNase P.[20] The type M RPRs are not active in the absence of their RPPs, however, thus ruling out the possibility of monitoring the gradual changes in RPR activity upon the addition of its RPPs. The topic of this thesis, the reconstitution of the RNase P holoenzyme from the archaeon *Methanobrevibacter smithii* (Msm) aims to fulfill that goal through the reconstitution of a mesophilic archaeal type A variant (Fig. 3).

In addition to the advantages listed above, a thorough study of *Msm* RNase P is expected to yield additional benefits. *Msm*, a methanogenic organism, has recently gained notoriety as one of the major archaeal representatives in the human gut.[32] Studies conducted in germ-free mice have found that co-colonization of the intestinal tract with *Bacteroides thetaiotaomicron*, a sugar-fermenting bacteria, and *Msm* enables the host organism to utilize a greater range of sugars than would otherwise be possible (Fig. 4).[33] The ability to extract additional calories from food leads to a greater tendency towards obesity, as shown in mouse models. The relative rarity of archaeal species within the human gut, when taken together with the difference in RNase P structure between the three domains of life and the essentiality of RNase P, suggests an attractive method to potentially curb an individual's *Msm* population. The ability to selectively inhibit archaeal RNase P based on the structure of the RPR and the interaction between RPR and RPPs would then provide a means to inhibit growth of the methanogen without affecting either the commensal gut bacteria or the animal host. Although the five archaeal RPPs share homology with the eukaryotic RPPs, the differences in structure and assembly of the holoenzyme may be sufficiently dissimilar to permit the selective targeting of the archaeal enzyme and not the eukaryotic variant. In fact, albeit not the most ideal proof-of-principle of this idea, we have recently designed cyclic peptide inhibitors of *Salmonella* RNase P that inhibit the growth of *Salmonella enterica* without affecting the growth of HeLa cells; these inhibitors are likely to function as peptidomimetics of the single bacterial RPP. Although we lack direct experimental evidence, it appears that these peptidomimics act by

competitively binding to the bacterial RPR and preventing association with its cognate RPP.[34] A similar method might be employed in the design of therapeutics against *Msm*.

The goals of this project are thus as follows:

1. To clone, overexpress, and purify the RPR and RPP subunits of *Msm* RNase P.
2. Reconstitute and biochemically characterize the activity of *Msm* RNase P.

MATERIALS AND METHODS

Cloning of RPR and RPP genes into expression vectors

Using the annotated *Msm* genome sequence, made available by the Washington University Center for Genome Sciences, the gene sequences of the RPR as well as the five RPPs (RPP21, RPP29, RPP30, POP5, and L7AE) were obtained (Table 1). The genes were then amplified by PCR using *Msm* genomic DNA as the template (courtesy of Dr. Jane Jackman, The Ohio State University). An *Eco*RI site was added to the reverse primer so as to place this restriction site immediately downstream to the coding sequence of the *Msm* RPR. Upon digestion of the PCR product with *Eco*RI, the RPR gene was cloned into pBT7,[35] which had been digested with *Stu*I and *Eco*RI. T4 DNA ligase (NEB) was used in all ligation reactions.

The PCR products containing the RPP21 and RPP30 open reading frames (ORFs) were amplified and digested with *Bam*HI, whose recognition site was incorporated immediately following the termination codon in the reverse primer. These *Bam*HI-digested PCR products were cloned into pET-15b (Novagen), which had been digested with *Nco*I, filled in with Klenow to create a blunt end, and subsequently digested with *Bam*HI.

The PCR products containing the POP5 and RPP29 ORFs were engineered with an *Nde*I recognition site encompassing the start codon in the forward primer, and a *Bam*HI recognition site immediately following the termination codon in the reverse primer. The PCR products of both genes were digested with *Nde*I and *Bam*HI and were cloned separately into pLANT-2b,[36] which had also been digested with *Nde*I and *Bam*HI.

The PCR product containing the L7Ae ORF was engineered with an *Nco*I recognition site encompassing the start codon in the forward primer, and a *Bam*HI recognition site immediately following the termination codon in the reverse primer. The PCR product was digested with both *Nco*I and *Bam*HI and cloned into pET-33b (Novagen), which had also been digested with *Nco*I and *Bam*HI.

Generation of the Msm RPR by in vitro *transcription*

To obtain the RPR transcript, the coding sequence was amplified by high-fidelity PCR with pBT7/*Msm* RPR as the template. This PCR product was digested with *Eco*RI and used as a template for T7 RNA polymerase-mediated run-off *in vitro* transcription. The reaction was extracted twice with a 25:24:1 mixture of phenol:chloroform:isoamyl alcohol, and then again with chloroform. The aqueous phase was dialyzed against double-distilled H_2O at 4°C, precipitated using two volumes of ethanol and one-tenth volume of 3 M NaOAc (pH 5.4), and the resulting pellet air-dried before re-suspension in water. The purified RPR was quantitated using an extinction coefficient of 2,517,100 M^{-1} cm^{-1} at 260 nm. The *Msm* RPR extinction coefficient was calculated by Integrated DNA Technologies (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/).

Overexpression and purification of RPP21●RPP29

To isolate the RPP21●RPP29 complex, pET-15b/RPP21 and pLANT-2b/RPP29 were co-transformed into *E. coli* BL21(DE3) cells to enhance the probability of their adopting a native fold, and to obtain a higher protein yield. The starter culture was grown in 1% (w/v) glucose in LB broth supplemented with 100 µg/mL carbenecillin and 35 µg/mL kanamycin. The culture was grown to an OD_{600} of ~0.6 prior to induction of overexpression using 1 mM IPTG and 1 mM $ZnCl_2$. The culture was then grown for an additional 17 h at 22°C. The cells were harvested and the cells from 125 mL of culture were re-suspended in 25 mL of a 1:3 mixture of buffers A [25 mM Tris-HCl (pH 7.5), 25 mM NaCl, 5 mM DTT, 0.1 mM PMSF] and B [25 mM Tris-HCl (pH 7.5), 2 M NaCl, 5 mM DTT, 0.1 mM PMSF], then sonicated on ice using a VibraCell™ sonicator (Sonics & Materials, amp 80, pulser 4, 1.5 min) four times, each sonication separated by a 2 min incubation on ice, and centrifuged (15 min, 9680 g, 4°C) to remove the cellular debris. The supernatant was removed and polyethylene imine (PEI) was added to a final concentration of

0.05% (v/v). The solution was incubated on ice for 30 min, and the resulting precipitate was centrifuged (15 min, 9680 g, 4°C) and discarded. Over the course of an hour, finely ground ammonium sulfate was added to the supernatant at 4°C with constant stirring to obtain a final concentration of 40% (w/v), and the resulting precipitate was removed through centrifugation (15 min, 9680 g, 4°C). Again, powdered ammonium sulfate was added at 4°C with stirring to a final concentration of 60% (w/v) over the course of an hour. The precipitate of proteins were removed by centrifugation (15 min, 9680 g), then resuspended in and dialyzed against buffer A at 4°C. The dialyzed solution was then filtered through a 0.45 µm membrane, loaded onto a 1-mL SP-Sepharose FF column (GE Healthcare) and eluted using a gradient of 25 mM to 2 M NaCl.

Overexpression and purification of RPP30 and POP5

E. coli BL21(DE3) cells were co-transformed with pLANT-2b/POP5 and pET-15b/RPP30 and protein overexpression was induced. However, upon induction (2 h at 37°C) very poor expression of POP5 resulted, most likely due to seven low-frequency usage codons contained within the first 15 codons of the POP5 transcript. (Studies on highly expressed *Eco* proteins during exponential growth have helped categorize a certain group as class II genes; for the purpose of this research, we define those codons whose usage is less than 10% of the total codon frequency in class II genes as low-frequency codons.) Attempting to use *E. coli* BL21(DE3) Rosetta cells (Novagen) to address the limitations from low-frequency codon usage did not produce significant improvements in POP5 overexpression. The Rosetta cells overexpress tRNA isoacceptors to remedy the low levels of select tRNA isoacceptors that decode Arg, Ile and Leu low-frequency codons. As *Msm* POP5 does not include many of these codons, it was not entirely unexpected that the Rosetta stain was unable to rescue the poor POP5 expression. Consequently, site-directed mutagenesis was used to introduce a set of silent mutations to replace six of the low-frequency codons with higher-frequency counterparts (Table 2). Due to the extensive mutagenesis of the

RPP that was required, the *Msm* POP5 gene was amplified in two parts and the two PCR products were then ligated, digested with *Nde*I and *Bam*HI, and re-cloned into pLANT-2b. Primers used for site-directed mutagenesis are shown below, with altered codons highlighted in bold font.

Forward: **CCG ACT CTG CGT** AAAAATAACAGGTATTTGGCATTGG

Reverse: **CGG CAG** AACCTTCAGCTTCATATGTATATCTCCTTCTTAAAGTTAAAC

E. coli BL21(DE3) cells were then co-transformed with the pET-15b/RPP30 and pLANT-2b/POP5 (codon-optimized) vectors, and overexpression of the two proteins was induced with 1 mM IPTG at an OD_{600} of ~0.6. The culture was then grown with shaking for 17 h at 22^{o}C. Cells from 125 mL of culture were harvested, re-suspended in 10 mL of a 1:1 mixture of buffers A and B, and sonicated. The purification of RPP30●POP5 followed a thematically similar approach to that of RPP21●RPP29 with the following modifications: upon resuspension of the 60% (w/v) ammonium sulfate-cut precipitate in buffer A, the solution was incubated on ice for 1 h. The resulting precipitate was removed by centrifugation (10 min, 18,000 g) and resuspended in 25mM Tris-HCl (pH 7.5), 2 M ammonium sulfate, 5 mM DTT, and 0.1 mM PMSF. The solution was filtered through a 0.45 µm membrane, and loaded onto a 1-mL HiTrap phenyl-Sepharose HP column (GE Healthcare), and eluted using a reverse gradient from 2 to 0 M ammonium sulfate.

Overexpression and purification of L7Ae

E. coli BL21(DE3) cells were transformed with pET-33b/L7Ae. Overexpression was induced with 1 mM IPTG and the culture was grown with shaking for an additional 3 h at 37^{o}C. After harvesting the cells by centrifugation, the cells from 125 mL of culture were re-suspended in 10 mL of a 1:1 mixture of buffers A and B and sonicated on ice using a VibraCell^{TM} sonicator (Sonics & Materials, amp 80, pulser 4, 1.5 min) four times, each sonication separated by a 2 min incubation on ice. After centrifuging the solution (15 min, 9680 g, 4^{o}C) to remove the cellular debris, PEI was added to a final concentration of 0.025% (v/v),

and the mixture was incubated on ice for 30 min. Ammonium sulfate was added to a final concentration of 80% (w/v) at 4°C with stirring, and the precipitate was pelleted and discarded. The supernatant was dialyzed against buffer A and filtered through a 0.45 µm membrane, then loaded onto a C4 reversed-phase HPLC column to further remove nucleic acid contamination. L7Ae was eluted using a gradient from 0.1% (v/v) trifluoroacetic acid (TFA) in water to 99.9% (v/v) acetonitrile and 0.1% (v/v) TFA. L7Ae-containing fractions were lyophilized and resuspended in 6 M guandinium-HCl, pH 6.5. The resuspended protein was dialyzed first against 1 L H_2O, then again directly into storage buffer [50 mM Tris-HCl (pH 8), 500 mM NH_4OAc].

By measuring the ultraviolet (UV) absorbance of the RPPs at 280 nm and using their respective extinction coefficients, the concentration of each RPP was determined. The quantitation of purified binary complexes based on their A_{280} values assumes that RPP21 and RPP29 (as well as RPP30 and POP5) are fully complexed in a 1:1 ratio. Absorbance measurements were also taken at 260 nm, and the 260/280 ratio was used to gain some insight into possible nucleic acid contamination.

Functional reconstitution assays

Immediately prior to each assay, 20 µM of the *Msm* RPR was folded in a thermocycler using the following program: 50 min at 50°C, 10 min at 37°C, 30 min at 37°C. During the second step, an equivalent volume of buffer [100 mM Tris-HCl (pH 7.5) 1 M NH_4OAc, and 15 mM $MgCl_2$] was added and the folded RPR was removed immediately following the third step and assayed within two hours of the folding event. Subsequent rounds of serial dilutions were performed using the assay buffer specified. Unless otherwise specified, these assays employed *Eco* pre-$tRNA^{Tyr}$ as the substrate; defined concentrations of unlabeled pre-$tRNA^{Tyr}$ were mixed with a trace amount internally labeled with α-^{32}P-GTP. Reactions were quenched using an equivalent volume of 8 M urea dye (8 M urea, 20% (v/v)

phenol, 0.05% (w/v) xylene cyanol, 0.05% (w/v) bromophenol blue, 1 mM EDTA) and products were electrophoresed on a denaturing 8% (w/v) polyacrylamide gel containing 7 M urea. Radioactive products were visualized using a Typhoon phosphorimager (GE Healthcare) and quantitated using ImageQuant software (GE Healthcare). While substrate incubated in the absence of RNase P served as the negative control, substrate cleaved by the *Eco* RNase P holoenzyme was employed as the positive control.

Reconstitution of the Msm RNase P holoenzyme

Assays were conducted in 50 mM Tris-HCl (pH 7.5), 500 mM NH_4OAc, and 7.5 mM $MgCl_2$. Fifty nM folded RPR and 500 nM RPPs (final concentrations) were incubated at 37^{o}C for 10 min to allow for assembly of the RNP. Substrate was then added to a final concentration of 250 nM, and the reaction was incubated at 37^{o}C for an additional 10 min. Similar reactions containing either the RPR alone or variable combinations of the RPPs were also incubated for 10 min.

RPR assay under single-turnover conditions

Assays were conducted in 50 mM Tris-HCl (pH 7.5), 2 M NH_4OAc, and a specified concentration of $MgCl_2$. Folded RPR (0.3 or 3 μM) was incubated at 37^{o}C for 14 h with 1.7 nM pre-tRNA^{Tyr}.

Michaelis-Menten analysis

Assays were conducted in 50 mM Tris-HCl (pH 8), 800 mM NH_4OAc, and 30 mM $MgCl_2$. Ten nM folded RPR and 100 nM RPPs (final concentrations) were pre-incubated at 37^{o}C for 10 min prior to addition of a defined concentration of substrate. Aliquots of this reaction incubated at 37^{o}C were withdrawn and quenched at defined time intervals (e.g. 3, 5, 7, and 9 min). The initial velocity of each

reaction was determined by plotting percent of product formed versus time, and calculating the initial slope. The initial velocities of the various reactions were then used to calculate the turnover numbers, and the resulting turnover numbers were plotted against the substrate concentration used in each reaction.

Reconstitution of partially assembled holoenzyme complexes

RPR + RPP21●RPP29: Assays were conducted in 50 mM Tris-HCl (pH 8), 100 mM NH_4OAc, and 120 mM $MgCl_2$. Two hundred and fifty nM RPR and 625 nM RPPs (final concentration) were incubated at 37°C for 10 min prior to addition of 10 µM of substrate. Aliquots were withdrawn after 1, 2, and 3 h and quenched with 8 M urea dye.

RPR + RPP30●POP5: Assays were conducted in 50 mM Tris-HCl (pH 8), 100 mM NH_4OAc, and 120 mM $MgCl_2$. Fifty nM RPR and 500 nM RPPs (final concentration) were pre-incubated at 37°C for 10 min prior to the addition of 10 µM substrate. Aliquots were withdrawn at 0.5, 1, and 2 h and quenched with 8 M urea dye, and the reactions were processed as described above.

RPR and *Pfu* L7Ae: Assays were conducted in 50 mM Tris-HCl (pH 8), 2 M NH_4OAc, and 600 mM $MgCl_2$. Two and one-half µM RPR and 5 µM *Pfu* L7Ae were pre-incubated at 37°C for 10 min prior to addition of 20 µM substrate. Aliquots were withdrawn at 1, 2, and 3 h and quenched with 8 M urea dye, and the reactions were processed as described above. *Pfu* L7Ae was a gift from Stella Lai, Laboratory of Prof. Mark Foster, Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH.

Determining the effects of Pfu L7Ae on the Mg^{2+} dependence of the Msm RNase P holoenzyme-catalyzed reaction

Assays were conducted in 50 mM Tris-HCl (pH 8), 800 mM NH_4OAc, and defined concentration of $MgCl_2$. Ten nM RPR and 100 nM RPPs (± *Pfu* L7Ae, final concentrations) were pre-incubated at 37°C for 10 min, then incubated for 6 min at 37°C after addition of 250 nM pre-$tRNA^{Tyr}$. The individual reactions contained 1, 2, 4, 6, 15 or 30 mM $MgCl_2$.

RESULTS AND DISCUSSION

Protein overexpression and purification

All five RPPs were successfully purified using the procedures outlined above (Figs. 5-7). During their respective overexpression and purification processes, RPP21●RPP29 and RPP30●POP5 were maintained as binary complexes, possibly enhancing protein overexpression, folding, and solubility. This also enabled us to attempt to quantitate the ratio of each protein within the complex using SYPRO® Ruby stain (Invitrogen; Fig. 8). Unlike Coomassie blue, which preferentially stains basic proteins, SYPRO® Ruby is thought to be an unselective protein staining method, rendering it quantitative upon correction for number of amide bonds. Three independent quantitaitons of the SYPRO® Ruby fluorescent signal gave an average RPP21:RPP29 ratio of 1:0.8 and an RPP30:POP5 ratio of 1:1.6. While we expected a 1:1 ratio, our inability to observe stoichiometric equivalence might be attributable to various reasons, including some selectivity exhibited by the SYPRO® Ruby dye.

Analysis of RPR activity in the absence of its cognate RPPs

The RPR was generated through T7 RNA polymerase-mediated run-off *in vitro* transcription (Fig. 9). One hallmark of type A RNase P holoenzymes is the ability of their RPR subunit to catalyze pre-tRNA cleavage in the absence of their RPP subunits. Consistent with this characterization, we found that the *Msm* RPR was active even in the absence of its RPPs under single-turnover conditions (Fig. 10).

Reconstitution of the Msm RNase P holoenzyme

We successfully reconstituted the *Msm* RNase P holoenzyme using the RPR and 4 RPPs (POP5●RPP30 + RPP21●RPP29). An initial reconstitution effort, even without optimization of assay conditions, resulted

in cleavage of pre-tRNA^{Tyr} with a turnover number of 0.065 min^{-1} (Fig. 11). A preliminary screen of Mg^{2+} concentrations using the 4 RPP-holoenzyme showed that 7.5 mM $MgCl_2$ is minimally required for realizing activity. Subsequent assays, conducted at 30 mM $MgCl_2$, resulted in a ten-fold higher turnover number, indicating that the optimization of mono- and divalent cations should be an important future goal for achieving maximum holoenzyme activity.

The activity of the 4 RPP-holoenzyme at different pre-tRNA^{Tyr} concentrations was subjected to Michaelis-Menten analysis and the data analyzed using KaleidaGraph (Synergy Software) (Fig. 12). The calculated k_{cat} is 1.10 ± 0.08 min^{-1} and K_m is approximately 280 ± 60 nM. In comparison, *Pfu* RNase P, another type A RNase P holoenzyme, has a k_{cat} of 9.5 min^{-1} and a K_m of 180 nM.[19] As mentioned previously, however, the *Pfu* RNase P holoenzyme originates from a thermophilic archaeon, and therefore its assays are conducted at 55^{o}C instead of 37^{o}C (the assay temperature utilized for *Msm* RNase P).

Msm L7Ae and heterologous reconstitution

The *Msm* RNase P holoenzyme was also assayed both with and without the fifth protein, L7Ae. From previous studies, L7Ae is known to affect the activity of the RPR+ 4 RPP-holoenzyme in several ways. L7Ae is capable of lowering the concentration of Mg^{2+} required for functional activity, increasing the k_{cat} and optimal temperature of activity, and decreasing the K_m for pre-tRNA cleavage.[37] It was found, however, that *Msm* L7Ae did not seem to affect the activity of the *Msm* RNase P holoenzyme in any of the aspects previously mentioned. We decided to further investigate this intriguing observation.

First, it is known that L7Ae is an RNA-binding protein, and a member of several different RNPs.[38,39] To remove nucleic acids with which L7Ae might be adventitiously associated during the purification process, it is frequently necessary to unfold and purify L7Ae on a C4 column using reversed-phase chromatography. Thus, the first possibility was that residual nucleic acid binding was interfering

with the ability of L7Ae to complex with the RPR and its cognate RPPs. The relatively high $A_{260/280}$ ratio (1.42) seemed to support this idea. Secondly, it was possible that *Msm* L7Ae might not be required for the function of the *Msm* RNase P holoenzyme at least under the conditions we employed.

To discriminate between these two hypotheses, we performed a heterologous reconstitution with *Pfu* L7Ae and *Msm* RPR + *Msm* POP5●RPP30+RPP21●RPP29. Despite the thermophilic nature of *Pfu* being somewhat incompatible with the mesophilic nature of *Msm*, sequence comparison of the *Pfu* and *Msm* L7Ae proteins showed high homology (Fig. 13). Indeed, a preliminary test for Mg^{2+} in the assay showed that while *Msm* L7Ae did not affect Mg^{2+} concentrations required for catalysis (neither the 4 RPP- nor the 5 RPP-holoenzyme showed activity at 5 mM $MgCl_2$), *Pfu* L7Ae was capable of activating the holoenzyme even at 2 mM $MgCl_2$ (Fig. 14). Thus it seems likely that the *Msm* L7Ae which we isolated and tested was inactive.

Partially reconstituted Msm RNase P

The complexes RPR + RPP21●RPP29, RPR + RPP30●POP5, and RPR + L7Ae (*Pfu*) were assembled as described in the methods section and assayed for activity in order to gain a clearer picture of the ability of the RPPs to enhance RPR activity. Although the RPR when added to L7Ae or RPP21●RPP29 or RPP30●POP5 showed no activity, the conditions tested mimicked those used with *Pfu* RNase P and thus might need to be empirically determined for *Msm* RNase P. Moreover, since the K_m of the *Msm* 4 RPP-holoenzyme is calculated to be higher than that of the *Pfu* counterpart (280 vs. 180 nM) and the k_{cat} was calculated to be somewhat lower (1.1 vs. 9.5 min^{-1}), it is possible that the incubation times we employed were also insufficient for detecting the activity of these partial complexes, which are likely to be weaker than the 4 RPP-containing holoenzyme.[21] In addition to optimizing multiple-turnover assay conditions, single-turnover assays also merit consideration before ruling out that the partial RNPs (RPR + L7Ae or RPR + RPP21●RPP29 or RPR + RPP30●POP5) are inactive. Barring activity from either single or

multiple-turnover conditions, it may be necessary to determine whether or not the RPPs complex to the RPR in the assay conditions employed through the use of techniques like gel mobility shift assays.

FUTURE DIRECTIONS

Results from the preliminary characterization of both the 4- and 5-RPP *Msm* RNase P holoenzymes, albeit encouraging, suggest that considerable optimization of assay conditions is required. Assays performed to date employed conditions found to be optimal for thermophilic counterparts of *Msm* RNase P. The underlying assumption with this approach was that both enzymes are sufficiently structurally and mechanistically similar to require similar mono- and divalent cation concentrations, although the Michaelis-Menten data obtained suggests that this assumption should be treated with caution, and that buffer conditions should be optimized specifically for the *Msm* enzyme before further biochemical studies are conducted. Moreover, the purification of functional *Msm* L7Ae is also an important pre-requisite to uncover its role in *Msm* RNase P catalysis.

Successful heterologous reconstitutions using *Msm* and *Pfu* RPPs identify opportunities for studying the basis of thermostability in a catalytic RNP such as RNase P. It is known that the RPR of thermophilic archaea tend to have much higher G-C content than those of their mesophilic counterparts (up to 70% G-C in the case of some hyperthermophiles). Whether or not there is a corresponding alteration in the structure or binding of the RPPs, however, is unknown. A heterologous reconstitution assay between *Msm* and *Pfu* was carried out (i.e., *Msm* RPR in conjunction with *Pfu* RPPs, and vice versa) as a method of testing the ability of the newly-reconstituted *Msm* RNase P holoenzyme to translate to the known thermophilic type A systems. While the *Pfu* RPR and *Msm* RPPs were capable of robust activity, the combination of *Msm* RPR and *Pfu* RPPs showed no activity (Fig. 15). While it is possible that the mesophilic RPR was not able to maintain an active conformation at the higher assay temperature (55^{o}C), it is also possible that the *Pfu* RPPs are unable to bind the *Msm* RPR (which has 60% A-U content).

Towards the ultimate goal of biochemical characterization through sm-FRET, more detailed studies of the contacts between the *Msm* RPR and its RPPs are also necessary. Current options for the

attachment of fluorophores tend to rely on the hybridization of fluorescently-labeled oligonucleotides to complementary sequences within the RNA. Certain regions vital for contacts with RPPs or for structural integrity, however, are not ideal locations for hybridization of a DNA oligonucleotide. Probing the structure for specific binding sites of the RPPs (e.g., through hydroxyl radical-mediated footprinting) can provide a general picture of the RPR sites that are accessible for hybridization and unlikely to interfere with RPP binding.

In addition to aiding sm-FRET studies, elucidation of structure-function relationships in *Msm* RNase P will also be important in the design of its potential inhibitors. The peptidomimetic proven to be selective against *Salmonella* RNase P mimics the RPR-binding portion of the bacterial protein subunit and consequently requires some prior knowledge of the general structure and interacting domains of the RPR and RPPs.[34] Such information on *Msm* RNase P will be vital before we can use it as a possible anti-obesity target.

In conclusion, *Msm* has considerable potential as a model for mesophilic archaeal type A RNase P. While it might be distinctive from the existing thermophilic RNase P variants with respect to substrate binding and catalysis, these differences might be adaptations necessary for its optimal function in a mesophilic environment. Experimental data have shown, however, that it is certainly capable of reconstituting with thermophilic type A RPPs, indicating that all of the structural elements necessary for RPR-RPP and RNP-substrate interaction within the thermophilic type A RNase P holoenzymes are likely to also be present in the mesophilic *Msm* model. Furthermore, the unique physiological niche of the *Msm* host organism lends an additional layer of interest to the fundamental question of RNase P evolution, structure-function, and molecular mechanism. It is also important to remember that archaeal RNase P ultimately serves as a model for eukaryotic nuclear RNase P, and so the ability to further characterize and dissect the activity of archaeal RNase P, regardless of the model system used, enables a better understanding of the catalytic RNP both horizontally and vertically.

REFERENCES

1. Greider CW, Blackburn EH. A Telomeric Sequence In The RNA of *Tetrahymena* Telomerase Required For Telomere Repeat Synthesis. (1989) *Nature*. **337**, 331-337.
2. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. The Structural Basis of Ribosome Activity In Peptide Bond Synthesis. (2000) *Science*. **289**, 920-930.
3. Valadkhan S, Manley JL. Splicing-Related Catalysis By Protein-Free snRNAs. (2001) *Nature*. **413**, 701-707.
4. Robertson HD, Altman S, Smith JD. Purification And Properties Of A Specific *Escherichia coli* Ribonuclease Which Cleaves A Tyrosine Transfer Ribonucleic Acid Precursor. (1972) *J. Biol. Chem.* **247**, 5243-5251.
5. Stark BC, Kole R, Bowman EJ, Altman S. Ribonuclease P: An Enzyme With An Essential RNA Component. (1978) *Proc. Natl. Acad. Sci.* **75**, 3717-3721.
6. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. The RNA Moiety of Ribonuclease P Is The Catalytic Subunit Of The Enzyme. (1983) *Cell.* **35**, 849-857.
7. Pannuci JA, Haas ES, Hall TA, Harris JK, Brown JW. RNase P RNAs From Some Archaea Are Catalytically Active. (1999) *Proc. Natl. Acad. Sci.* **96**, 7803-7808.
8. Kikovska E, Svard SG, Kirsebom LA. Eukaryotic RNase P RNA Mediates Cleavage In The Absence Of Protein. (2007) *Proc. Natl. Acad. Sci.* **104**, 2062-2067.
9. Hall TA, Brown JW. The Ribonuclease P Family. (2001) *Methods Enzymol.* **341**, 56-77.
10. Holzmann J, Frank P, Loffler E, Bennett KL, Gerner C, Rossmanith W. RNase P Without RNA: Identification And Functional Reconstitution Of The Human Mitochondrial tRNA Processing Enzyme. (2008) *Cell.* **135**, 462-474.
11. Wang MJ, Davis NW, Gegenheimer P. Novel Mechanisms For Maturation Of Chloroplast Transfer RNA Precursors. (1988) *EMBO J.* **7**, 1567-1574.
12. Green CJ, Rivera-Leon R, Void BS. The Catalytic Core Of RNase P. (1996) *Nucleic. Acids Res.* **24**, 1497-1503.
13. Loria A, Pan T. Domain Structure Of The Ribozyme From Eubacterial RNase P. (1996) *RNA.* **2**, 551-563.
14. Lai LB, Vioque A, Kirsebom LA, Gopalan V. Unexpected Diversity Of RNase P, An Ancient tRNA Processing Enzyme: Challenges And Prospects. (2010) *FEBS Letters*. **584**, 287-296.
15. Walker SC, Engelke DR. Ribonuclease P: The Evolution Of An Ancient RNA Enzyme. (2006) *Crit. Rev. Biochem. Mol.* **47**, 77-102.
16. Evans D, Marquez SM, Pace NR. RNase P: Interface Of The RNA And Protein Worlds. (2006) *Trends Biochem. Sci.* **31**, 333-341.
17. Gopalan V. Uniformity Amid Diversity In RNase P. (2007) *Proc. Natl. Acad. Sci.* **104**, 2031-2032.
18. Hall TA, Brown JW. Archaeal RNase P Has Multiple Protein Subunits Homologous To Eukaryotic Nuclear RNase P Proteins. (2002) *RNA.* **8**, 296-306.
19. Fukuhara H, Kifusa M, Watanabe M, Terada A, Honda T, Numata T, Kakuta Y, Kimura M. A Fifth Protein Subunit Ph1496p Elevates The Optimum Temperature For Ribonuclease P Activity From *Pyrococcus horikoshii* OT3. (2006) *BBRC.* **343**, 956-964.
20. Cho IM, Lai LB, Susanti D, Mukhopadhyay B, Gopalan V. Ribosomal Protein L7Ae Is A Subunit Of Archaeal RNase P. (2010) *Proc. Natl. Acad. Sci.* **107**, 14573-14578.
21. Tsai HY, Pulukkunat DK, Woznik WK, Gopalan V. Functional Reconstitution And Characterization Of *Pyrococcus furiosus* RNase P. (2006) *Proc. Natl. Acad. Sci.* **103**, 16147-16152.
22. Kouzuma Y, Mizoguchi M, Takagi H, Fukuhara H, Tsukamoto M, Numata T, Kimura M. Reconstitution Of Archaeal Ribonuclease P From RNA And Four Protein Components. (2003) *BBRC.* **306**, 666-673.
23. Pulukkunat DK, Gopalan V. Studies On *Methanocaldococcus jannaschii* RNase P Reveal Insights Into The Roles Of RNA And Protein Cofactors In RNase P Catalysis. (2008) *Nucleic. Acids Res.* **36**, 4172-4180.

24. Li D, Willkomm DK, Hartmann RK. Minor Changes Largely Restore Catalytic Activity Of Archaeal RNase P RNA From *Methanothermobacter thermoautotrophicus*. (2009) *Nucleic. Acids Res.* **37**, 231-242.
25. Xu Y, Amero CD, Pulukkunat DK, Gopalan V, Foster MP. Solution Structure Of An Archaeal RNase P Binary Protein Complex: Formation of the 30-kDa Complex between *Pyrococcus furiosus* RPP21 and RPP29 Is Accompanied by Coupled Protein Folding and Highlights Critical Features for Protein-Protein and Protein-RNA Interactions. (2009) *J. Mol. Biol.* **393**, 1043-1055.
26. Kazantsev AV, Rambo RP, Karimpour S, Stalucia J Jr, Tainer JA, Pace NR. Solution Structure Of RNase P RNA. (2011) *RNA.* **17**, 1159-1171.
27. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragon A. Crystal Structure Of The RNA Component of Bacterial Ribonuclease P. (2010) *Nature*. **468**, 784-751.
28. Chen WY, Pulukkunat DK, Cho IM, Tsai HY, Gopalan V. Dissecting Functional Cooperation Among Protein Subunits In Archaeal RNase P, A Catalytic Ribonucleoprotein Complex. (2010) *Nucleic Acids Res.* **38**, 8316-8327.
29. Stryer L, Haugland RP. Energy Transfer: A Spectroscopic Ruler. (1967) *Proc. Natl. Acad. Sci.* **2**, 719-726.
30. Klostermeier D, Millar DP. Time-Resolved Fluorescence Energy Transfer: A Versatile Tool For The Analysis Of Nucleic Acids. (2001) *Biopolymers.* **61**, 159-179.
31. Ha T, Ogletree DF, Chemla DS, Selvin PR, Weiss S. Probing The Interaction Between Two Single Molecules: Fluorescence Resonance Energy Transfer Between A Single Donor And A Single Acceptor. (1996) *Proc. Natl. Acad. Sci.* **13**, 6264-6268.
32. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA. Diversity Of The Human Intestinal Microbial Flora. (2005) *Science.* **208**, 1635-1638.
33. Samuel BS, Gordon JI. A Humanized Gnotobiotic Mouse Model Of Host-Archaeal-Bacterial Mutualism. (2006) *Proc. Natl. Acad. Sci.* **103**, 10011-10016.
34. Go C. (2010) M.S. Thesis, Department of Biochemistry, The Ohio State University, Columbus, OH.
35. Tsai HY, Lai LB, Gopalan V. A Modified pBluescript-Based Vector For Facile Cloning And Transcription Of RNAs. (2002) *Anal. Biochem*. **303**, 214-217.
36. Finklestein J, Antony E, Hingorani MM, O'Donnell M. Overproduction And Analysis Of Eukaryotic Multiprotein Complexes In *Escherichia coli* Using A Dual-Vector Strategy. (2003) *Anal. Biochem*. **319**, 78-87.
37. Chen WY, Gopalan V. Unpublished data
38. Klein DJ, Schmeing TM, Moore PB, Steitz TA. The Kink-Turn: A New RNA Secondary Structure Motif. (2001) *EMBO J.* **20**, 4214-4221.
39. Charron C, Manival X, Clery A, Senty-Segault V, Charpentier B, Marmier-Gourrier N, Branlant C, Aubry A. The Archaeal sRNA Binding Protein L7Ae Has A 3D Structure Very Similar To That Of Its Eukaryal Counterpart While Having A Broader RNA-binding Specificity. (2004) *J. Mol. Biol.* **342**, 757-773.
40. Walker SC, Engelke DR. A Protein-Only RNase P In Human Mitochondria. (2008) *Cell*. **135**, 412-414.

FIGURES

Figure 1. RNP subunit composition of the RNase P holoenzyme across the three domains of life. Figure reproduced from ref. 40.

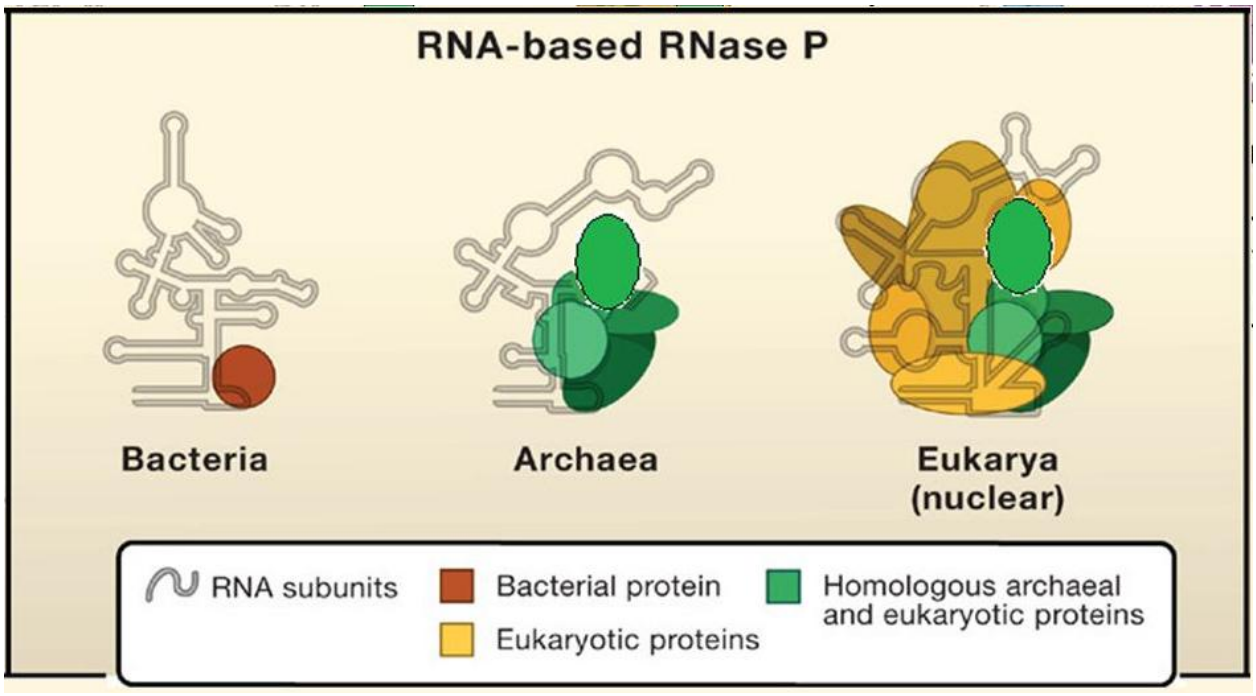

Figure 2. Secondary structure models of RPRs from three domains of life.

Paired regions are indicated by P1, P2, etc., and psuedoknots and distal contacts are indicated by solid lines. Conserved nucleotides are shown in bold, uppercase letters. Reproduced from ref. 17.

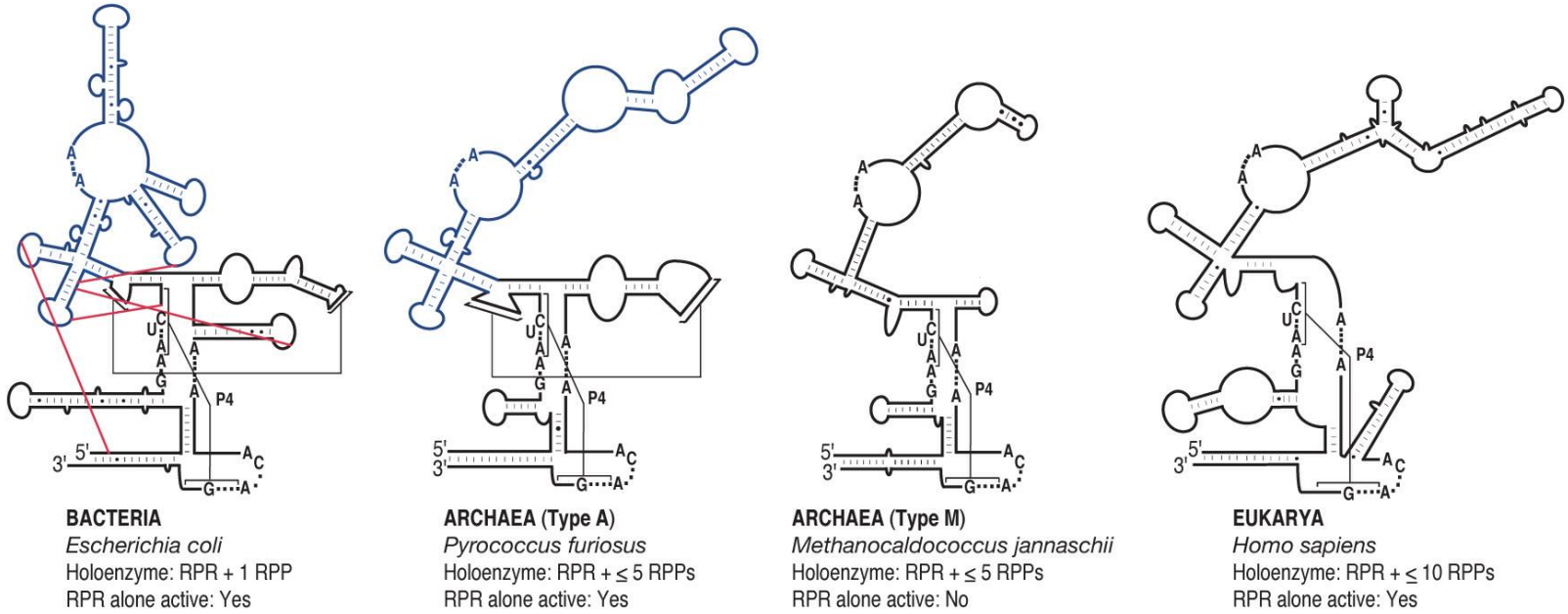

Figure 3. Secondary structure representation of the *Msm* RPR.

Figure provided courtesy of Dr. Lien B. Lai (Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH).

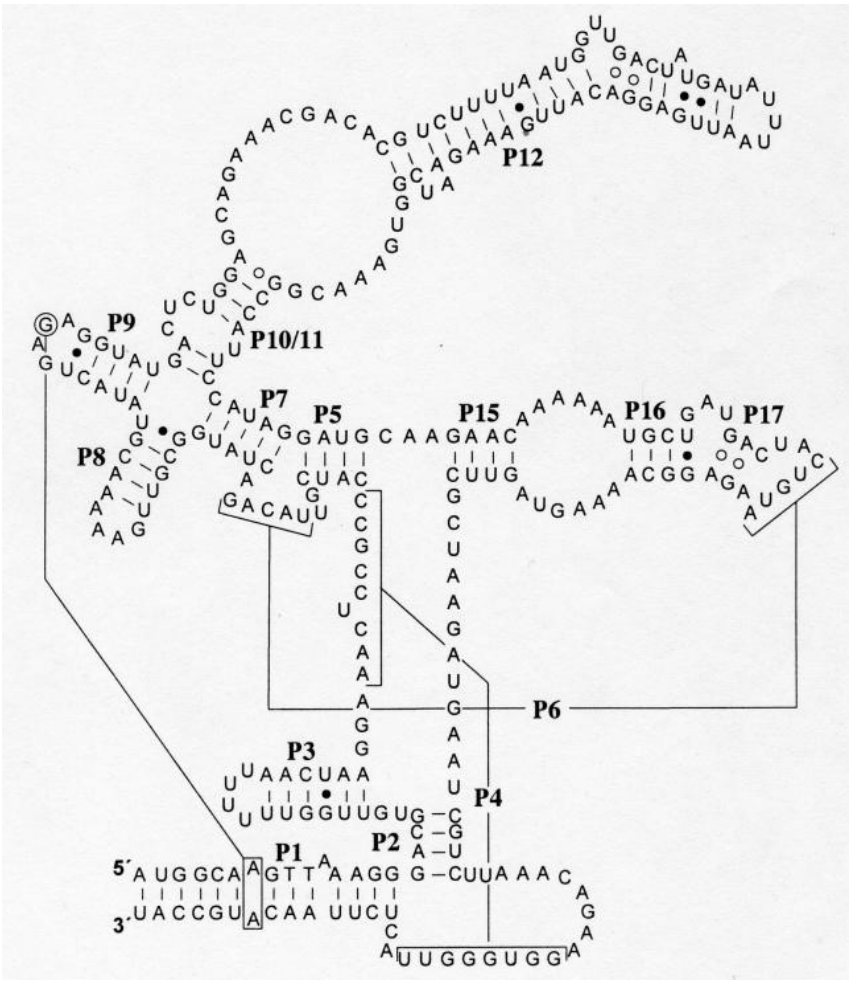

Figure 4. Effects of gut microbiome composition on intestinal contents.

Cecal sugar content of germ-free mice upon colonization/co-colonization with the indicated species. GF = germ free, Bt = *Bacteroides thetaiotaomicron*, Ms = *Methanobrevibacter smithii*. Figure reproduced from ref 33.

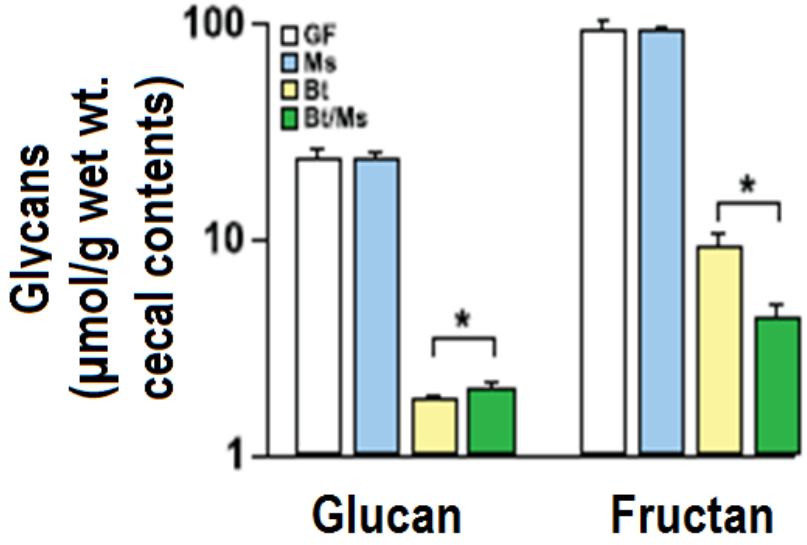

Figure 5. Analysis of *Msm* RPP21●RPP29 co-overexpression.

Induction and co-overexpression of *Msm* RPP21●RPP29 in *E. coli* BL21(DE3) cells. Analysis performed using 15% (w/v) polyacrylamide-SDS gel electrophoresis (tricine). The gel was stained with Coomassie blue. M = protein ladder, UN = uninduced control, IN = induced culture.

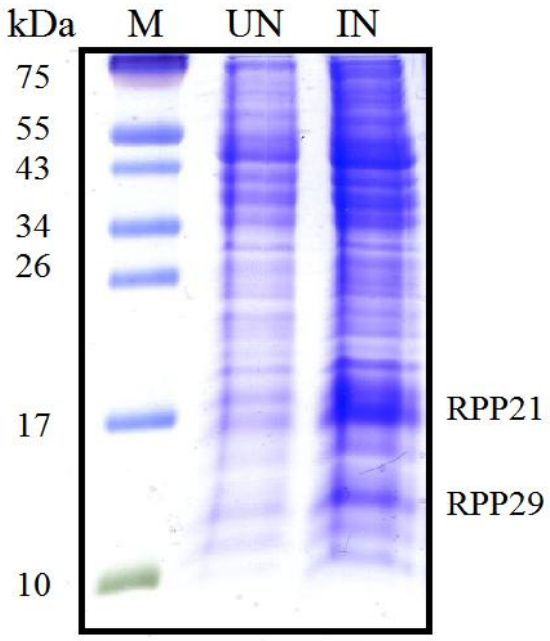

Figure 6. Analysis of *Msm* RPP30●POP5 co-overexpression

Induction and co-overexpression of *Msm* RPP30●POP5 (post-mutagenesis) in *E. coli* BL21(DE3) cells. Analysis performed using 15% (w/v) polyacrylamide-SDS gel electrophoresis (tricine). The gel was stained with Coomassie blue. M = protein ladder, UN = uninduced control, IN = induced culture.

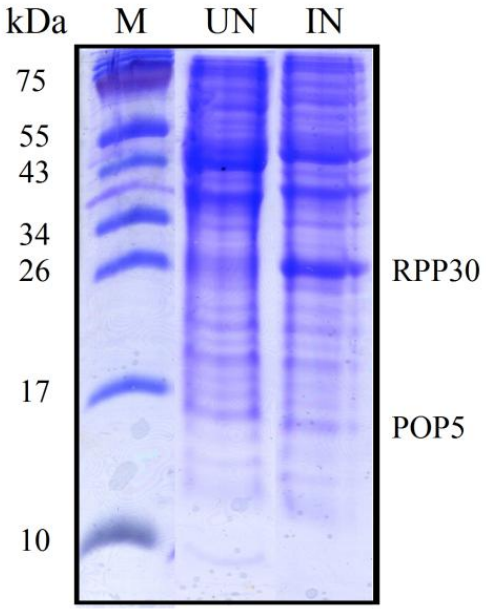

Figure 7. Analysis of *Msm* L7Ae induction

Induction and overexpression of *Msm* L7Ae in *E. coli* BL21(DE3) cells. Analysis performed using 15% (w/v) polyacrylamide-SDS gel electrophoresis. The gel was stained with Coomassie blue. M = marker, UN = uninduced control, IN = induced culture.

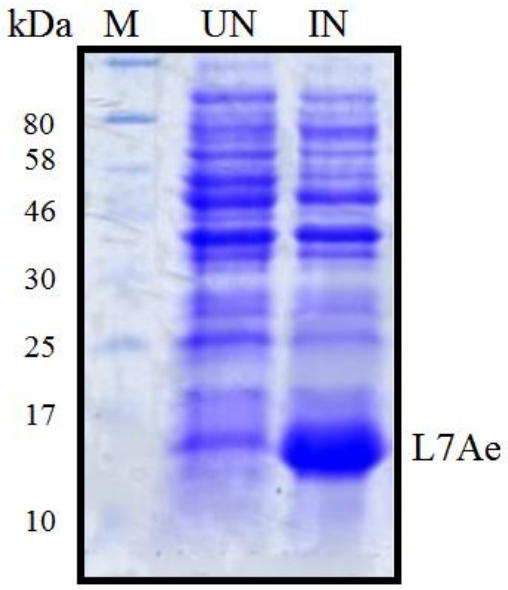

Figure 8. Quantitation of the stoichiometry of RPPs in binary complexes.

Analysis performed using 15% (w/v) polyacrylamide-SDS gel electrophoresis (tricine). The gel was stained with SYPRO® Ruby (Invitrogen). M = protein ladder. Quantitation of fluorescence performed using Typhoon phosphorimager (GE Healthcare; Fluorescence detection, 610 BP emission filter). The molar ratio of RPP21:RPP29 was determined to 1:0.8 while that of RPP30:POP5 ratio was 1:1.6.

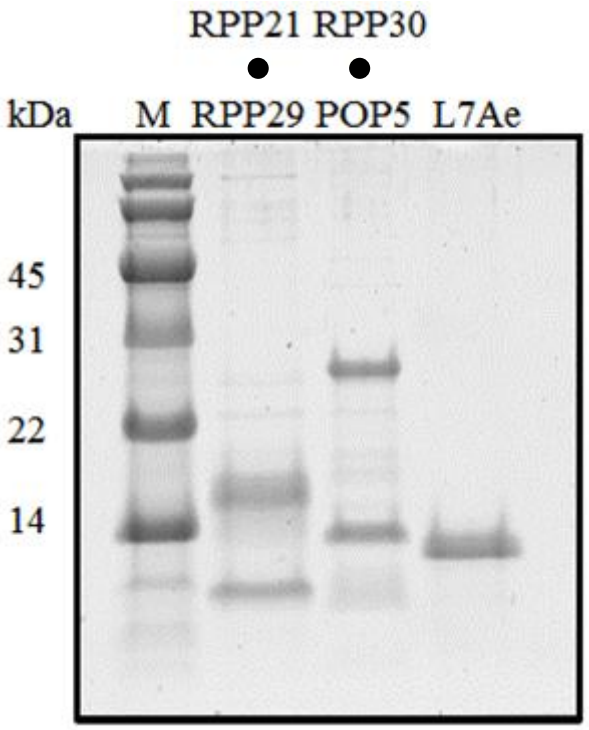

Figure 9. Analysis of *in vitro*-transcribed *Msm* RPR.

Agarose [1.5% (w/v)] gel electrophoresis to assess the size (287 nts) and integrity of the *in vitro*-transcribed *Msm* RPR. The gel was stained with ethidium bromide.

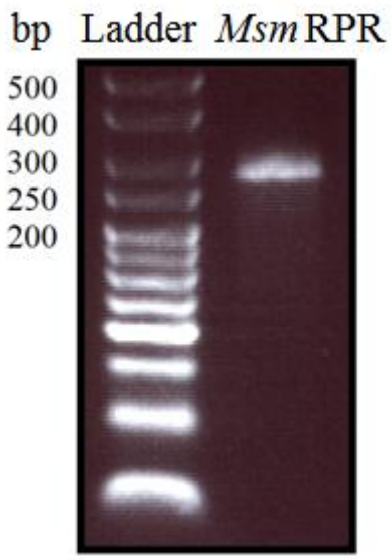

Figure 10. Assay of the *Msm* RPR alone under single-turnover conditions.

Msm RPR was incubated with 2.6 nM pre-tRNA^{Tyr} for 14 h at 37^{o}C.
PC = positive control (generated form processing of pre-tRNA^{Tyr} by *in vitro* reconstituted *Eco* RNase P), NC = negative control (substrate incubated without any enzyme). Concentrations of RPR and $MgCl_2$ used in each reaction are as shown.

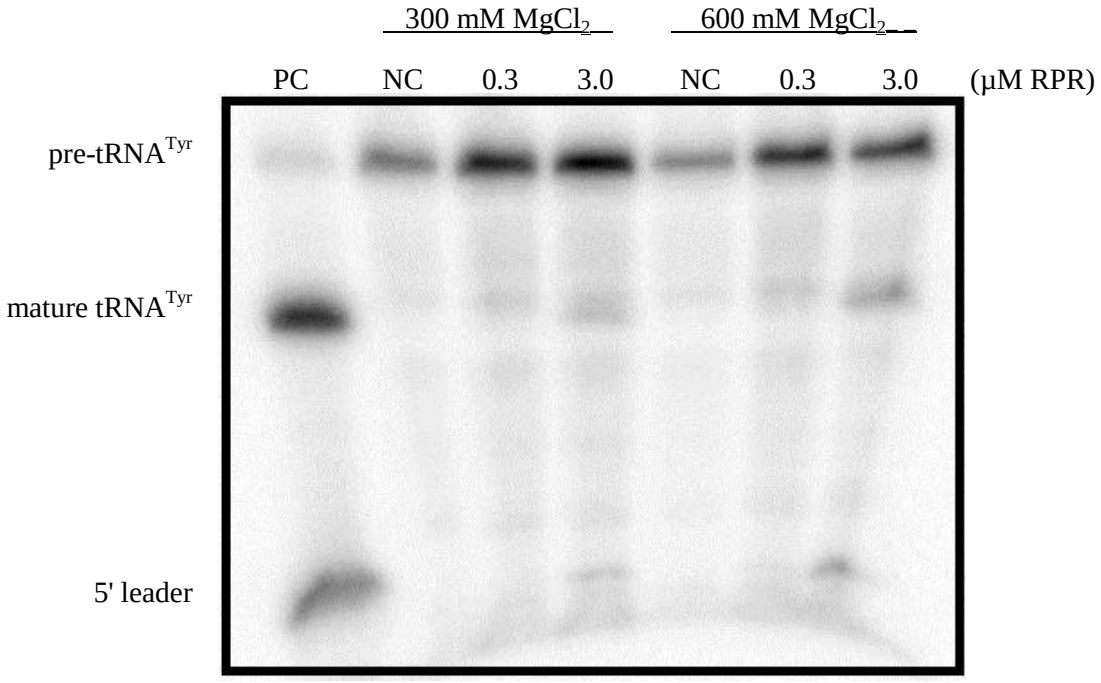

Figure 11. Functional reconstitution of the *Msm* RNase P holoenzyme.

Msm RPR (50 nM) was incubated with a 10-fold excess of RPPs and 250 nM pre-tRNATyr for 10 min at 37°C. PC = positive control (generated by pre-tRNA^{Tyr} processing by *in vitro* reconstituted *Eco* RNase P), NC = negative control (substrate incubated without any enzyme).

Components	PC	NC	1	2	3	4
RPR	-	-	+	-	+	+
RPP21●RPP29	-	-	-	+	+	+
RPP30●POP5	-	-	-	+	+	+
L7Ae	-	-	-	+	-	+

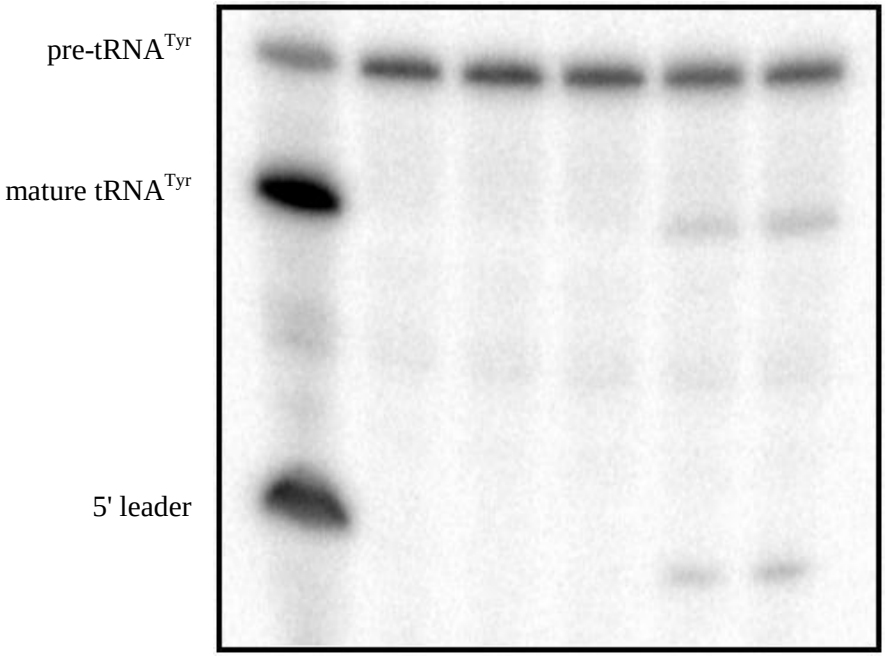

Figure 12. Michaelis-Menten analysis of the pre-tRNA^{Tyr} cleavage by the *Msm* RNase P holoenzyme.

Msm RNase P holoenzyme assembled from the RPR and 4 RPPs (RPP21●RPP29 + RPP30●POP5). Data analysis was performed using KaleidaGraph (Synergy Software).

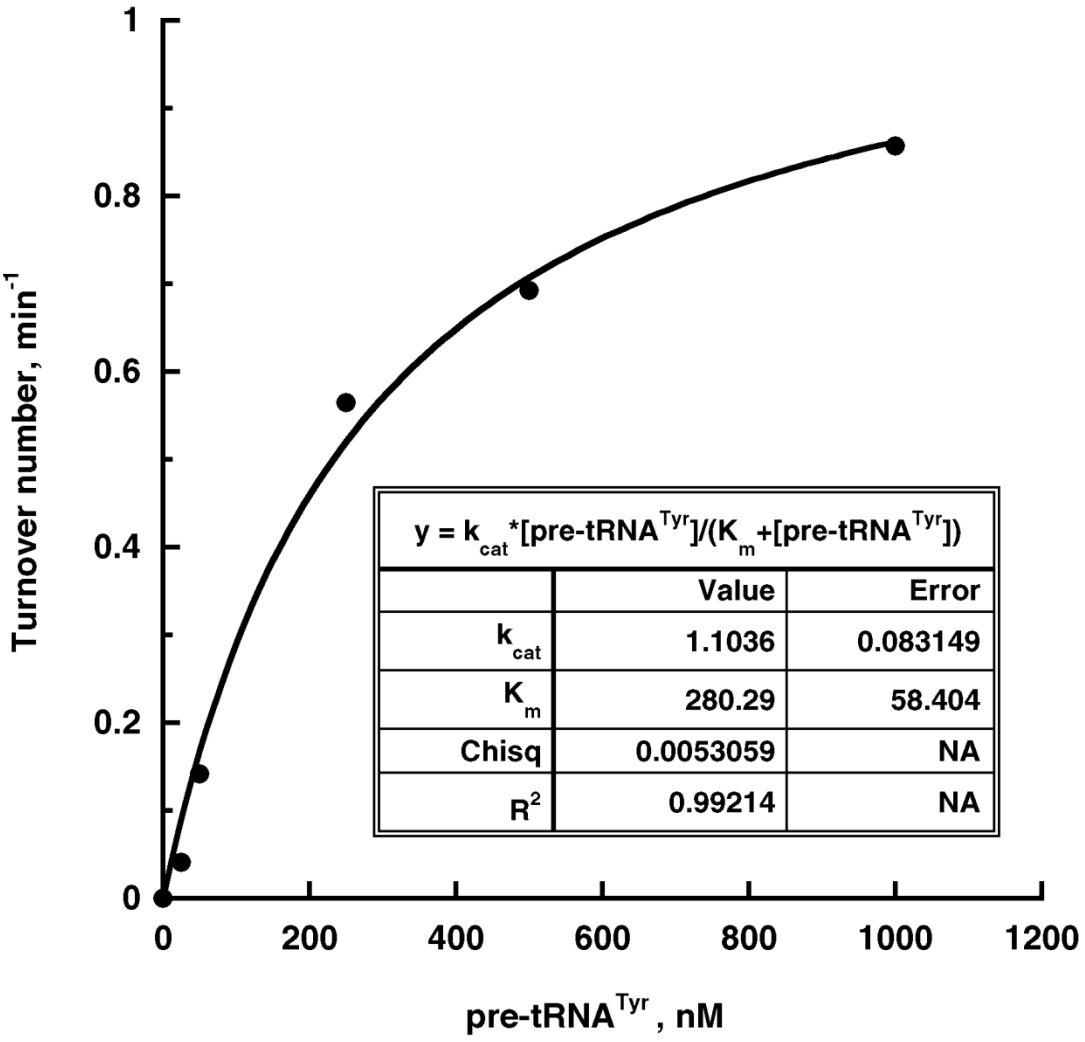

Figure 13. Alignment of *Pfu* and *Msm* L7Ae protein sequences.

Alignment performed by NCBI protein BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Query = *Pfu* L7Ae, Sbjct = *Msm* L7Ae.

```
Score = 130 bits (328), Expect = 1e-32, Method: Compositional matrix adjust.
Identities = 70/121 (58%), Positives = 92/121 (76%), Gaps = 0/121 (0%)

Query  1   MAKAIYVKFDTPEELANQAEEALKTAQDSGKVAKGTNEVTKFIERGDAALVVIAEDVDPA
           MAK  YVKF+ P+ELA +A +A++ A+D+GK+ KGTNE TK +ERG A LV+IAEDVDP
Sbjct  2   MAKPSYVKFEVPKELAEKALQAVEIARDTGKIRKGTNETTKAVERGQAKLVIIAEDVDPE

Query  61  EIVAHIPVLADEKEIPYIYLPTKEQVGGAAGLTVGTASACIVDAGDAEGDVAEIVEKIAE
           EIVAH+P L +EKEIPYIY+P+K+++G AAG+ V  AS  I++ G A   V EI  K+ E
Sbjct  62  EIVAHLPPLCEEKEIPYIYVPSKKELGAAAGIEVAAASVAIIEPGKARDLVEEIAMKVKE

Query  121  L  121
             L
Sbjct  122  L  122
```

Figure 14. Heterologous reconstitution using *Msm* and *Pfu* RNase P subunits.

Reconstitution of *Msm* RNase P (RPR + RPP21●RPP29 + RPP30●POP5) with *Pfu* L7Ae. *Msm* RPR (10 nM) was incubated with a 10-fold excess of RPPs (± *Pfu* L7Ae) for 9 min at 37°C with 2 mM $MgCl_2$. PC = positive control (generated by pre-$tRNA^{Tyr}$ processing by *in vitro* reconstituted *Eco* RNase P), NC = negative control (substrate incubated without any enzyme).

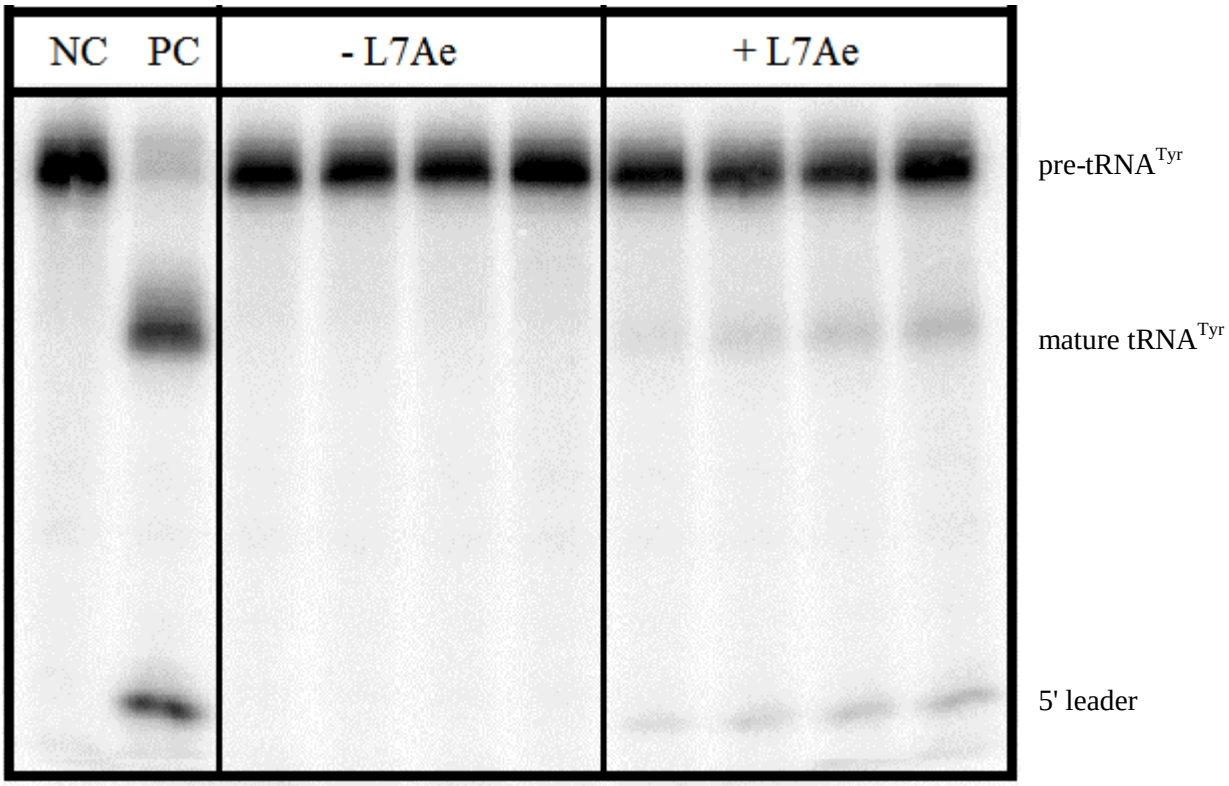

Figure 15. Heterologous reconstitution of *Msm* and *Pfu* RPR and RPPs.

Msm or *Pfu* RPR (50 nM) was incubated with a 10-fold excess of RPPs and 250 nM pre-tRNA^{Tyr} for 10 min at the indicated temperatures. PC = positive control (generated by pre-tRNA^{Tyr} processing by *in vitro* reconstituted *Eco* RNase P), NC = negative control (substrate incubated without any enzyme). MM = homologously reconstituted *Msm* RNase P holoenzyme, PM = *Pfu* RPR + *Msm* RPPs, MP = *Msm* RPR + *Pfu* RPPs, PP = homologously reconstituted *Pfu* RNase P holoenzyme.

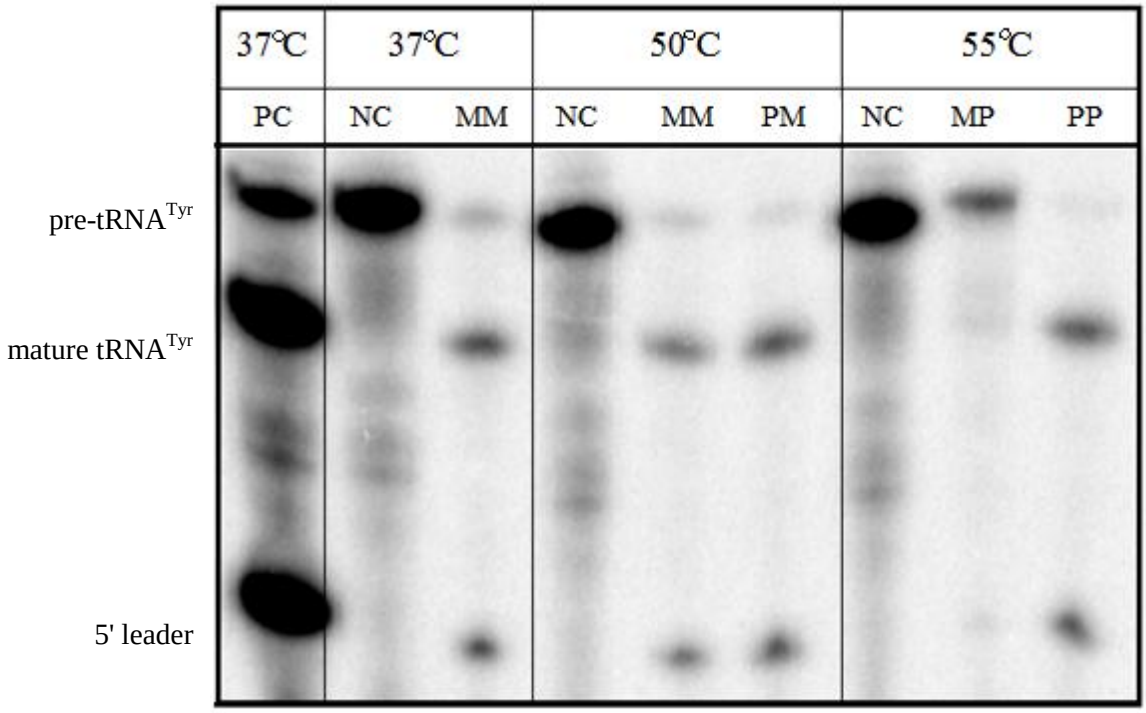

Table 1. Characteristics of *Msm* RPPs.

Data obtained from the respective protein sequences and calculated by the Expert Protein Analysis System proteomics server, maintained by the Swiss Institute of Bioinformatics (http://ca.expasy.org/tools/#proteome).

Protein	Molecular Weight (kDa)	Theoretical pI
RPP21	14.7	10
RPP29	10.4	9
RPP30	26.8	8.4
POP5	14	9.4
L7AE	12.9	4.3

Table 2. Analysis of codon usage in *Msm* POP5.

Identity of the codon, its encoded amino acid, and the frequency of usage in *Eco* class II genes are included for each position. Post-mutagenesis sequence and resulting frequency of usage in *Eco* class II genes are also given for each position. Low frequency usage codons are shown in bold font.

Msm POP5 ORF					
Codon position	Codon	Amino acid	*Eco* usage (Class II genes)	Mut. Codon	*Eco* usage (Class II genes)
1	AUG	Met	1.00	AUG	1.00
2	AAG	Lys	0.21	AAG	0.21
3	**CUU**	**Leu**	**0.06**	**CUG**	**0.77**
4	AAG	Lys	0.21	AAG	0.21
5	GUU	Val	0.40	GUU	0.40
6	**CUU**	**Leu**	**0.06**	**CUG**	**0.77**
7	**CCC**	**Pro**	**0.02**	**CCG**	**0.72**
8	CCU	Pro	0.11	CCG	0.72
9	**ACA**	**Thr**	**0.05**	**ACU**	**0.29**
10	**CUU**	**Leu**	**0.06**	**CUG**	**0.77**
11	**AGG**	**Arg**	**0.003**	**CGU**	**0.64**
12	AAA	Lys	0.79	AAA	0.79
13	AAU	Asn	0.17	AAU	0.17
14	AAC	Asn	0.83	AAC	0.83
15	**AGG**	**Arg**	**0.003**	**AAG**	**0.003**